Open

ORIGINAL ARTICLE

Disambiguating ventral striatum fMRI-related bold signal during reward prediction in schizophrenia

RW Morris[1,2], A Vercammen[1,2,3], R Lenroot[1,2,3], L Moore[1,3], JM Langton[4], B Short[2,3,5], J Kulkarni[6], J Curtis[5], M O'Donnell[5], CS Weickert[1,2,3] and TW Weickert[1,2,3]

[1]Schizophrenia Research Institute, Darlinghurst, NSW, Australia; [2]School of Psychiatry, University of NSW (UNSW), Kensington, NSW, Australia; [3]Neuroscience Research Australia, Randwick, NSW, Australia; [4]School of Psychology, UNSW, Kensington, NSW, Australia; [5]Kiloh Centre, Prince of Wales Hospital, Randwick, NSW, Australia and [6]Alfred Psychiatric Research Centre, Melbourne, VIC, Australia

Reward detection, surprise detection and prediction-error signaling have all been proposed as roles for the ventral striatum (vStr). Previous neuroimaging studies of striatal function in schizophrenia have found attenuated neural responses to reward-related prediction errors; however, as prediction errors represent a discrepancy in mesolimbic neural activity between expected and actual events, it is critical to examine responses to both expected and unexpected rewards (URs) in conjunction with expected and UR omissions in order to clarify the nature of ventral striatal dysfunction in schizophrenia. In the present study, healthy adults and people with schizophrenia were tested with a reward-related prediction-error task during functional magnetic resonance imaging to determine whether schizophrenia is associated with altered neural responses in the vStr to rewards, surprise prediction errors or all three factors. In healthy adults, we found neural responses in the vStr were correlated more specifically with prediction errors than to surprising events or reward stimuli alone. People with schizophrenia did not display the normal differential activation between expected and URs, which was partially due to exaggerated ventral striatal responses to expected rewards (right vStr) but also included blunted responses to unexpected outcomes (left vStr). This finding shows that neural responses, which typically are elicited by surprise, can also occur to well-predicted events in schizophrenia and identifies aberrant activity in the vStr as a key node of dysfunction in the neural circuitry used to differentiate expected and unexpected feedback in schizophrenia.
Molecular Psychiatry (2012) **17**, 280–289; doi:10.1038/mp.2011.75; published online 28 June 2011

Keywords: associative learning; prediction error; reward; schizophrenia; ventral striatum

Introduction

We develop and maintain expectations about real world contingencies based on past experiences. It is precisely when those expectations are not met, termed 'prediction errors', that we learn and update our expectations. People with schizophrenia have difficulty in learning from prediction errors,[1–3] which may be related to ventral striatum (vStr) dysfunction.[4,5] Understanding the neuronal basis for this difficulty in schizophrenia may help to develop treatments that could restore normal learning.

Three distinct hypotheses regarding the functions of the vStr have been proposed, which can be distinguished according to the effects reward and surprise have in four different conditions of reinforcement: unexpected reward (UR), expected reward (ER), unexpected reward omission (UO) and expected reward omission (EO). In the reward hypothesis, the vStr is activated in response to reward or reinforcement (see Figure 1a) and pathophysiology leads to diminished reward-related responding and an inability to experience and learn from reward, for example, anhedonia.[6,7] A second hypothesis, referred to here as the surprise hypothesis, suggests that the vStr responds to unpredicted or surprising events. Evidence from healthy adults shows the vStr is activated by unpredictable rewards relative to predictable rewards.[8,9] As such, the vStr may differentiate surprising from predicted outcomes, including URs and UOs (see Figure 1b), and lack of differential responding may lead to inappropriate interest or attention directed to neutral or predictable stimuli.[10,11] In the third hypothesis, the prediction-error hypothesis, the vStr is activated and deactivated in response to the mismatch between actual and predicted events consistent with formal models of associative learning.[4,5,12] This hypothesis suggests the neural response to URs and UOs will be diametrically opposed and distinguishable from the effect of
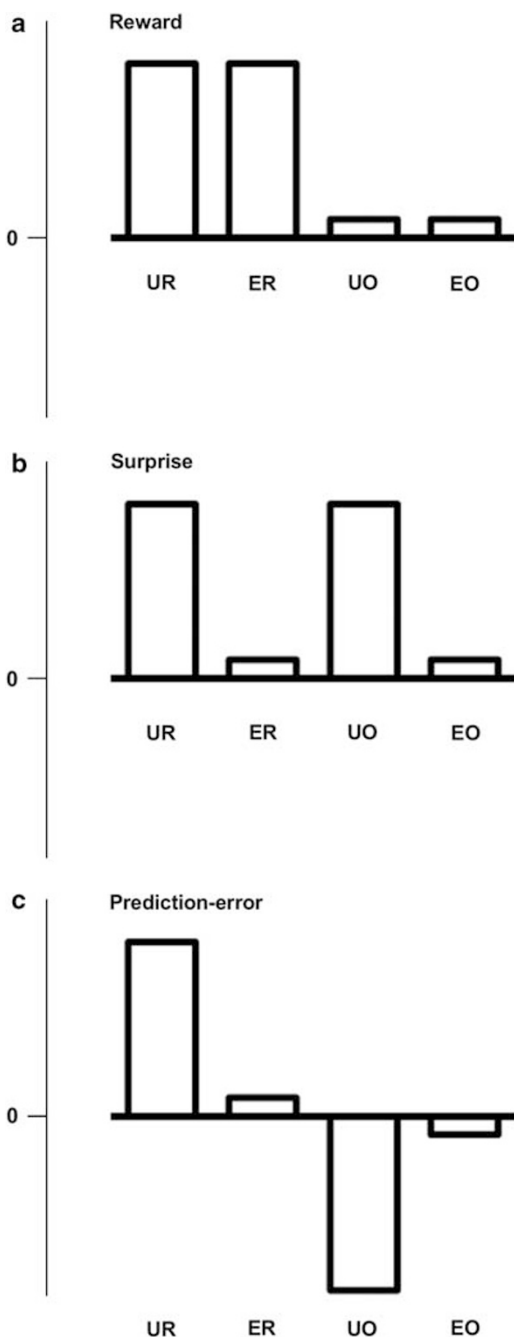
Figure 1 Three models of possible activation across four conditions of reward (UR: unexpected reward, ER: expected reward, UO: unexpected omission of reward, EO: expected omission of reward). Panel **a** indicates a main effect of reward; panel **b** indicates a main effect of surprise; panel **c** indicates an interaction in which surprise modulates the response to reward in a bidirectional manner.

We tested whether neural activity patterns in the midbrain and striatum of people with schizophrenia relative to healthy adults were consistent with reward, surprise or prediction error using functional magnetic resonance imaging (fMRI) and a reward-related prediction error task.

Schizophrenia is associated with aberrant responses to primary and secondary rewards in the vStr.[15,16] The neural response to both expected and unexpected primary rewards (juice) is diminished during a Pavlovian conditioning task.[17] The response to secondary reinforcers may also be aberrant in schizophrenia when the reward is unexpected.[18] Murray et al.[19] examined reward-related prediction-error signals in first-episode psychosis patients, some of whom were antipsychotic naive, and found that these patients displayed attenuated differential midbrain activity during unexpected rewarded trials relative to ERs; however, this attenuated difference may result from greater activation during the comparison condition, that is, during ERs. For instance, the response to an ER can appear exaggerated in the vStr of chronically ill people with schizophrenia.[20] Thus, the exact nature of the abnormal ventral striatal reward-related prediction-error signal in schizophrenia is currently uncertain and may be due to decreased neural responses to UR or increased neural responses to ER.

No study to date has separately measured and reported the neural response to expected and UR (and UO and EO) in people with schizophrenia to determine whether deficits are specific to unexpected or expected conditions. Identifying whether deficits are specific to URs or ERs (or both) is needed to clarify the nature of the aberrant prediction-error signal in people with schizophrenia and treat problems with feedback learning driven by prediction errors. Measuring the effect of reward, surprise and the interaction between both (that is, prediction errors) will provide a better understanding of striatal dysfunction in schizophrenia and potentially provide a neural platform to test the effectiveness of new treatments.

Materials and methods

Participants
Sixteen healthy adults and 21 chronically ill people with schizophrenia or schizoaffective disorder were recruited for this study. Five patients were excluded for excessive head motion (>2 mm), inadequate task performance (non-responding) or structural abnormalities leaving 16 people with schizophrenia, all of whom were receiving second-generation antipsychotic medication. Trained clinicians provided Structured Clinical Interview for the Diagnostic and Statistical Manual diagnosis,[21] intelligence quotient estimates and reading scores as an estimate of premorbid intelligence quotient in schizophrenia,[22,23] and symptom severity ratings using the Positive and Negative Syndrome Scale (PANSS).[24] The Edinburgh handedness inventory was used to measure handed-

surprise or reward alone (see Figure 1c).[13] In the prediction-error hypothesis, vStr dysfunction may result in aberrant neural responses when no mismatch between expectations and outcomes has occurred,[14] or logically, this hypothesis could also predict the aberrant absence of response when a mismatch has occurred (that is, a failure to respond).

ness in each participant. All volunteers provided written informed consent and this study was approved by the South East Sydney and Illawarra Area Health Service and the University of New South Wales Human Research Ethics committees. See the Participants section in the Supplementary information for further details.

fMRI acquisition

We acquired 968 whole brain T2* weighted echoplanar images; slice thickness: 3 mm, 1 mm gap, 31 axial slices in ascending order, repetition time: 2000 ms, echo time: 30 ms, flip angle: 90 degrees, matrix: 112×112, field of view: 240 mm using a Phillips Achieva 3T scanner at NeuRA, Randwick, NSW. A T1-weighted high-resolution anatomical scan and MPRAGE were obtained for each participant for registration and screening; repetition time: 5.4 ms, echo time: 2.4 ms, field of view: 256 mm, matrix: 256×256, sagittal plane, slice thickness: 1 mm with no gap, 180 slices.

Task design

Participants were asked to 'pretend to play a card game'. A series of Pavlovian cue-outcome association trials was presented in which participants were instructed to predict a reward (an image of nine $50 Australian notes), which was contingent upon presentation of one (the regularly winning 'trump' card) of four distinct playing cards, see Task Design in Figure 2. To establish expectations between the trump card and the money stimulus, each participant was initially trained to six consecutive correct responses before entering the scanner. During the scan session, 120 trials were presented in which 96 trials were consistent with training and 24 trials were inconsistent with training (catch trials), in order to ensure prediction errors occurred. The first 12 trials in the scanner were consistent with training to initially reinforce learned expectations from training. After the first 12 trials, a partial reinforcement schedule was used (P (reward 'trump' card) = 0.78] to ensure learned expectations were maintained. The trump card and reward stimulus were presented on half the trials in the entire session, with the following expected frequencies for each condition: 0.4 ERs; 0.4 EO; 0.1 URs; and 0.1 UO. Any prediction incongruent with the outcome was included as an unexpected event, thus the obtained number of trials in each condition differed for each participant according to their idiosyncratic response history. The interval between trials was jittered and ranged between 4051–6040 ms, during which a crosshair was presented. The identity of the trump card was counterbalanced between diamonds, squares, circles and triangles. Participants were required to respond 2–4 s before presentation of the reward stimulus. Failure to respond resulted in a missed response recorded and these trials were ignored in the analysis. Failure to respond to more than 25% of trials in any one condition resulted in exclusion as a non-responder.
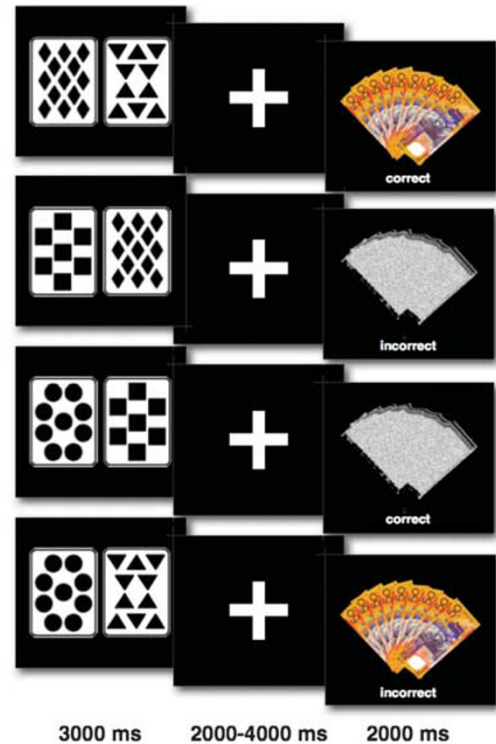

Figure 2 Examples of card cues and outcomes trial types in prediction task. Four different card stimuli were used (diamonds, triangles, squares and circles) in which one card was trump (diamonds in this example). Money (reward stimulus) was either presented or omitted. Predictions were either correct or incorrect. Top row: expected reward; second row: unexpected reward omission; third row: expected reward omission; fourth row: unexpected reward. Stimulus and feedback duration are provided below.

For more detailed methods, see Supplementary information (Methods).

Data analysis

The percentage of correct trials in the first 12 trials of the card game during the scan, previous to the introduction of catch trials, was used to confirm the success of training. Differences between groups in the percentage of correct trials across the entire scan (excluding catch trials) were tested with independent, two-tailed t-tests.

Image processing and analysis were performed using SPM5. Images for each participant were realigned to the first image in the sequence to correct for head motion, and slice time correction was applied to adjust for differences in the time of slice acquisition. Images were spatially normalized to the Montreal Neurological Institute (MNI) template (ICBM452) using a non-linear 12-parameter affine transformation. All images were smoothed to minimize noise and residual differences in gyral anatomy with a Guassian kernel set at 8 mm full-width at half-maximum. Data sets were also manually screened for scan stability (< 2 mm head movement), image

artifacts, structural abnormalities and successful normalization to ensure image quality was equivalent between groups.

The first-level (individual-subject) analysis modeled events as stick functions. Eight statistical regressors were defined based on the time series of stimulus onsets and the time series of each participant's responses, as measured by the Presentation™ software: four regressors of interest representing reward outcomes correctly and incorrectly predicted (ER and UR); and non-reward outcomes correctly and incorrectly predicted (EO and UO); and four regressors of no interest including two cue regressors, representing card presentations predicted by the participant to result in reward and non-reward images (reward cards and non-reward cards) and two response regressors for reward and non-reward predictions (reward responses and non-reward responses). Six regressors from the motion correction procedure were included in the event-related design to account for slight variation due to head movement. Each of the eight task regressors was convolved with the canonical hemodynamic response function, and individual *t*-contrasts of the four outcome regressors were used to generate contrast images that would be taken to a second-level analysis. As we were interested in voxels deactivated to UOs (negative prediction errors), we calculated the negative parameter estimates (regressor coefficients) of the UO and EO regressors to generate contrast images for these conditions.

Whole brain
Second-level (group) analyses were run separately in the healthy participants and people with schizophrenia. A three-way analysis of group, reward and surprise directly compared both groups. The contrast images from the healthy adults of the four types of outcome events (UR, ER, EO and UO) were included in a 2×2 factorial ANOVA, where reward was the first factor (reward versus omission) and surprise was the second factor (unexpected versus expected). The main effect of reward, inclusively masked with a directional *t*-contrast (UR + ER > UO + EO, see Figure 1a), was tested to determine whether presenting an image of nine 50 dollar bills evoked responses in reward-related regions of the brain. We used an inclusive mask in each main effect and interaction test to restrict the significant voxels to those that show the relevant predicted direction of effect shown in Figure 1. The main effect of surprise was tested to reveal regions that responded to unexpected events over expected events in an unidirectional fashion (UR + UO > ER + EO, Figure 1b). The interaction between surprise and reward (UR > ER; EO > UO, Figure 1c) was tested to identify regions activated to URs relative to ERs (that is, positive prediction-error signals) and deactivated to UOs relative to EOs (that is, negative prediction-error signals) to reveal regions displaying a prediction-error signal. We report significant clusters and their peak voxels that survive the false-discovery rate (a correction for multiple comparisons) $P < 0.05$ in the whole brain.[25,26]

Regions of interest (ROI)
We also performed a ROI analysis incorporating the midbrain and striatum because these regions have been implicated in reward and prediction-error processing[9,27] and are also implicated in the pathology of schizophrenia.[2,6,11,17,28–31] The ROI was defined using the PickAtlas tool (Wake Forest University, Winston-Salem, NC, USA)[32] including the midbrain, caudate and putamen regions based on the MNI template (Colin27) included as a single ROI. A 2×2 factorial ANOVA testing the main effects and the interaction in each group, as well as a three-way ANOVA testing the group comparisons, was performed in the same manner as the whole-brain analysis described above. We report any significant voxels that survive false-discovery rate $P < 0.05$. The location of significant peak voxels in the ROI was confirmed by reference to the Atlas of the Human Brain.[33] The vStr was defined as a region extending in stereotaxic coordinates: 2–18 mm rostral (*y*-axis); 5 mm dorsal and −5 mm ventral (*z*-axis); and ± 5 to 15 mm lateral (*x*-axis).[33] We also report the parameter estimates ('beta weights') at a subset of peak voxels to confirm the direction of effect at each location and the contribution of each condition (UR, UO, ER and EO) to the interaction effect. Inspection of the parameter estimates allowed us to assess whether the apparent neural activity represented a prediction-error signal according to the three criteria:[13] (1) significant positive activation after URs relative to ER; (2) significant deactivation after UO relative to EO and (3) no significant change from baseline after ERs and EO.

Covariate-of-interest analyses
We performed a covariate analysis in SPM5 (false-discovery rate $P < 0.05$) to determine whether aberrant neural activity in the ROI of people with schizophrenia was associated with symptom severity (positive, negative and total PANSS score) or antipsychotic dosage in chlorpromazine equivalents.[34,35] Correlations between each symptom score and the parameter estimates (beta weights) from the ER and EO conditions within the ROI were calculated.

Results

Participants
Demographic results are shown in Table 1. As expected, education levels and current intelligence quotient estimates were significantly greater in healthy adults relative to people with schizophrenia. However, there was not a significant difference in WTAR scores between groups and the WTAR score was above average (that is, 100) in schizophrenia, suggesting that our patients were not disadvantaged before the onset of schizophrenia. The mean total PANSS score in our sample of people with

Table 1 Demographic data for healthy adults and people with schizophrenia

	Patients (16)	Controls (16)	t (df 30)	P-value
Age	33.0	32.9	0.82	0.75
Females	7	8		
Years of education	13.8 (1.9)	16.4 (1.8)	4.47	< 0.01
WAIS-III IQ estimate	99.9 (13.9)	121.4 (19.2)	3.36	< 0.01
WTAR reading score	108.4 (9.7)	114.6 (5.4)	1.92	0.06
Edinburgh handedness score	80 (5.5)	84 (6.0)	0.95	0.35
PANSS score (total)	68.0 (20.6)			
PANSS score (positive)	16.4 (4.6)			
PANSS score (negative)	13.9 (7.0)			
Medication	Clozapine (6)			
	Risperidone (4)			
	Olanzapine (4)			
	Quetiapine (1)			
	Amisulpride (1)			

Abbreviations: PANSS, Positive and Negative Syndrome Scale; IQ, intelligence quotient; WAIS-III, Weschler Adult Intelligence Scale 3rd Edition; WTAR, Weschler Test of Adult Reading.

schizophrenia was in the mild to moderately impaired range with respect to illness severity.

Behavior
Both groups successfully learned to predict the winning outcome before the introduction of catch trials, the mean (s.e.m.) percentage of correct trials during the first 12 trials of the scan were 92 (3) and 88 (4) for HA and SC groups, respectively, $t(30) = 0.89$, $P = 0.38$. Both group performed at similar levels throughout the entire scan: the percentage of correct trials (excluding the 24 catch trials) for each group was 90 (3) and 85 (4) for HA and SC groups, respectively, $t(30) = 1.17$, $P = 0.25$ (see Supplementary Figure S1 and the Supplementary Results page 3 for further details).

Imaging
The whole-brain analysis of healthy adults found a significant main effect of reward throughout the insula, the left and right inferior frontal gyrus, as well as the cingulate and the midbrain, however there was no significant main effect of surprise. The test of the critical interaction effect (UR > ER; EO > UO) in healthy adults revealed significant clusters in the insula, frontal gyrus, anterior cingulate, caudate (vStr) and midbrain; however, examination of the parameter estimates revealed only clusters in the vStr met all three criteria for prediction-error signals[13] (see Supplementary Figure S2 and Supplementary Results, page 4). The vStr was further analyzed in the ROI analysis described below. A post-hoc test for group differences in the effect of URs versus ERs (Group-(UR > ER)) found the response to URs among people with schizophrenia was blunted in the parietal cortex, insula and anterior cingulate (see Supplementary Figure S3 for cingulate results) while the response to ERs was exaggerated in the caudate and vStr of people with schizophrenia. No significant group differences were revealed by the main effect of reward

or surprise in the whole-brain analysis. See the Imaging Results in the Supplementary information and Supplementary Tables S1, S2 and S3 for the complete list of whole-brain results.

Healthy adults: ROI analysis
Table 2 provides the complete list of peak voxels for each significant contrast tested in the ROI analysis of healthy adults, people with schizophrenia and the group comparison. The main effect of reward found the reward stimulus (money image) elicited the largest cluster of activation as well as the global peak voxel in the midbrain (Healthy adults: UR + ER > UO + EO). The main effect of Surprise (UR + UO > ER + EO) did not reveal any significant voxels within the ROI, indicating that unexpected events (that is, URs and UOs combined) did not produce significantly more or less neural activation than expected events. The critical interaction contrast (UR > ER; EO > UO) to reveal prediction-error signals in the healthy adults ROI found significant peak voxels throughout the left and right caudate with the largest cluster of significant voxels in the midbrain. The largest peak activation occurred in the left vStr with significant activation also in the right vStr (Supplementary Figure S4). Inspection of the parameter estimates confirmed neural activation peaks in the vStr and the ventral midbrain (in or near the substantia nigra/ventral tegmental area) occurred in a manner consistent with prediction-error signals.[13] A post-hoc analysis comparing activation during URs versus ERs without the reward omission conditions (healthy adults: UR > ER) revealed significant activity throughout the caudate, putamen and midbrain in similar regions to the Surprise × Reward interaction contrast.

People with schizophrenia: ROI analysis
The main effect of reward (UR + ER > UO + EO) found significant activation in the midbrain in people with

Table 2 ROI results in healthy adults and people with schizophrenia

	T (voxel)	Z	k	FDR	MNI (x, y, z)	Label
Healthy adults						
Main effect of surprise (UR + UO ≠ ER + EO)				*No significant voxels (FDR > 0.05)*		
Main effect of reward	7.57	6.83	1351	< 0.01	(6, −23, −13)	Midbrain (ventral)
(UR + ER > UO + EO)	6.09	4.99		< 0.01	(6, −23, −9)	Midbrain
	4.75	3.75		< 0.01	(−4, −29, −19)	Midbrain
	5.19	4.73	29	< 0.01	(28, 20, −2)	Putamen
Reward × surprise interaction	6.81	5.88	525	< 0.01	(−8, 8, −2)	Caudate (vStr)
(UR > ER; EO > UO)	6.20	5.47		< 0.01	(−10, 4, 2)	Putamen (vStr)
	5.77	5.16		< 0.01	(−10, 2, 8)	Caudate (vStr)
	3.21	3.08		0.01	(−20, 20, 2)	Putamen
	6.16	5.84	253	< 0.01	(10, 8, 4)	Caudate (vStr)
	6.16	5.43	109	< 0.01	(28, 20, −2)	Putamen
	2.53	2.46		0.04	(18, 20, 0)	Caudate
	5.96	5.30	1422	< 0.01	(8, −18, −10)	Midbrain
	5.86	5.22		< 0.01	(−4, −22, −18)	Midbrain
	5.47	4.94		< 0.01	(−8, −20, −14)	Parahippocampal
	5.31	4.82		< 0.01	(4, −32, −22)	Midbrain
	3.86	3.64		0.01	(−6, −16, −4)	Midbrain
Simple effect of unexpected reward	4.93	4.5	147	< 0.01	(12, 6, 10)	Caudate
(UR > ER)	4.87	4.46	817	< 0.01	(4, −24, −10)	Midbrain
	4.83	4.42		< 0.01	(−2, −24, −18)	Midbrain
	4.25	3.96		< 0.01	(8, −16, −10)	Midbrain
	4.44	4.11	37	< 0.01	(28, 20, −2)	Putamen
	4.43	4.1	178	< 0.01	(−10, 10, 2)	Caudate (vStr)
	2.7	2.61		0.03	(−16, 20, 2)	Caudate
People with schizophrenia						
Main effect of surprise (UR + ER < > UO + EO)				*No significant voxels (FDR > 0.05)*		
Main effect of reward	6.86	5.96	1181	< 0.01	(6, −25, −6)	Midbrain
(UR + ER > UO + EO)	5.17	4.73		< 0.01	(−6, −28, −11)	Midbrain (ventral)
	4.25	3.99		0.01	(2, −24, −18)	Midbrain (ventral)
Reward × surprise interaction	4.48	4.18	69	0.02	(−2, −28, −2)	Midbrain
(UR > ER; EO > UO)	3.56	3.39		0.04	(8, −32, −10)	Midbrain
Simple effect of unexpected reward (UR > ER)				*No significant voxels (FDR > 0.05)*		
Group comparisons: (HA > SC)						
Group × surprise				*No significant voxels (FDR > 0.05)*		
Group × reward				*No significant voxels (FDR > 0.05)*		
Group × reward × surprise	4.56	4.37	257	0.02	(−8, 10, 0)	Caudate (vStr)
	4.12	3.98	303	0.02	(8, 8, 0)	Caudate (vStr)
	3.83	3.71		0.03	(10, 16, −2)	Caudate (vStr)
Group × simple effect of UR	4.06	3.92	1054	0.02	(−4, 8, −2)	Caudate (vStr)
Group × (UR > ER)	3.86	3.74		0.03	(8, 8, −2)	Caudate (vStr)
	3.29	3.21		0.04	(10, 8, 10)	Caudate
	3.59	3.5	617	0.04	(4, −18, −20)	Midbrain
	3.59	3.49		0.04	(4, −24, −10)	Midbrain

Abbreviations: EO, expected reward omission; ER, expected reward; HA, health adults; MNI, Montreal Neurological Institute; ROI, regions of interest, SC, people with schizophrenia; UO, unexpected reward omission; UR, unexpected reward; vStr, ventral striatum. FDR < 0.05.

schizophrenia. We did not detect any voxels showing a main effect of Surprise (UR + UO > ER + EO) within the ROI in people with schizophrenia. The critical interaction (UR > ER; EO > UO) to identify prediction-error signals revealed a small cluster of voxels in the midbrain. However, inspection of the parameter

estimates at peak voxels in the midbrain revealed ERs elicited some activation above baseline and URs did not produce activity that was significantly greater than baseline, thus this region failed to meet the criteria for a prediction-error signal[13] in schizophrenia. A post-hoc analysis examining the simple effect of URs (people with schizophrenia: UR > ER) failed to reveal any significant clusters or voxels within the ROI.

Group comparison: ROI analysis
Testing for group differences in the main effect of reward (Controls > Patients in UR + ER > UO + EO) or the main effect of Surprise (Controls > Patients in UR + UO > ER + EO) did not reveal any significant voxels in the ROI analysis. The three-way inter-action testing for aberrant prediction-error-related activation among people with schizophrenia relative to healthy adults confirmed significant diagnostic differences existed in the vStr (Table 2: Group × Re-ward × Surprise). Two clusters, in the left (Figure 3a) and right (Figure 3b) vStr, were significantly different between healthy adults and people with schizophrenia. In the left vStr, examination of the parameter estimates confirmed a normal prediction-error signal occurred in the healthy adults, but activation among people with schizophrenia to URs and UOs was severely attenuated (Figure 3a, right panel). Activation after ER in the left vStr of people with schizophrenia was significantly greater than baseline, though not significantly greater than healthy adults.

In the right vStr of people with schizophrenia, activation and deactivation after ER and EOs was significantly different than baseline and higher than healthy adults (Figure 3b, right panel). Thus, responses in the left and right vStr did not fit the criteria for a prediction-error signal in people with schizophrenia. Finally, testing the post-hoc contrast for group differences in the simple effect of URs versus ERs (Group × (UR > ER)) found significant differences in the caudate, putamen, midbrain and the vStr, a region already identified by the Group-Reward × Surprise contrast.

Covariate-of-interest analyses
Greater BOLD activation in the left and right vStr during the ER condition was correlated with higher total PANSS scores (see Supplementary Figure S5A and B). Significant correlations also existed between negative symptoms (PANSS subscale) and aberrant neural activity during ERs in the right vStr, $r = 0.71$, $P < 0.05$, and within the left vStr, $r = 0.68$, $P < 0.05$. This is consistent with other reports that the severity of negative symptoms is related to aberrant ventral striatal activity in people with schizophrenia.[28,29,36] There were no significant correlations with positive symptom scores within the ROI (r's < 0.5). Greater BOLD deactivation in the right vStr during ER omissions was correlated with higher antipsychotic (CPZ) doses (Supplementary Figure S5c).
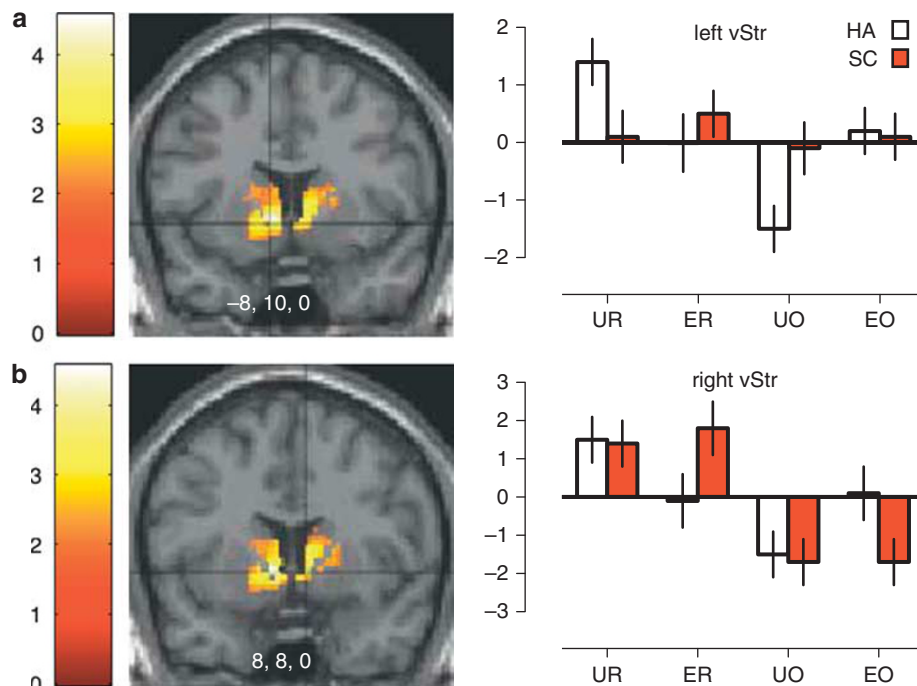

Figure 3 Regions of significant group differences in prediction-error signals in the ROI among healthy adults and people with schizophrenia. Color bars represent the *t*-value of the three-way interaction (false-discovery rate $P < 0.05$). Parameter estimates for each condition are shown on the right. (**a**) In the left vStr, people with schizophrenia displayed a relatively blunted level of BOLD activity during unexpected outcomes (UR and UO conditions). (**b**) In the right vStr, people with schizophrenia displayed a relatively exaggerated level of BOLD activity during expected outcomes (ER and EO conditions).

Discussion

We found bilateral neuronal responses in the vStr were largely driven by prediction errors in healthy adults: URs produced activation while UOs produced deactivation. Among people with schizophrenia, neural responses to unexpected outcomes in the left vStr were severely attenuated while neural responses to ERs were exaggerated bilaterally in the vStr. Our results show that vStr activity in people with schizophrenia does not correspond to the normal activity pattern during prediction-error detection and thus may fail to properly encode expected and unexpected outcomes in schizophrenia.

In healthy adults, we found neural responses consistent with prediction-error signals in the vStr; conversely, the neural responses tested outside this region did not meet criteria for a prediction-error signal, which is consistent with other reports in healthy adults of prediction-error signals restricted to the vStr.[37] Healthy adults had a bidirectional activation pattern in the vStr[38–40] perhaps reflecting the activity of direct projections from midbrain dopamine neurons, which increase and decrease neural response during positive and negative prediction errors.[41] Some previous fMRI studies have found a unidirectional, rather than a bidirectional, effect of surprise in the vStr[8] and concluded that phasic dopamine release in the vStr is valence independent (and thus related to salience). However, these studies omit the reward omission condition, and when reward omission was removed as a condition in the analysis, we found that the midbrain, insula, cingulate and parietal cortex were activated by URs in healthy adults and people with schizophrenia. However, there were no regions unidirectionally activated by unexpected events across conditions of reward and non-reward (that is, no significant main effect of Surprise). Other studies that included reward omissions have also found a bidirectional effect (both activations and deactivations) of surprise consistent with our results rather than a unidirectional effect of surprise.[27,42,43] In general, the BOLD deactivations observed after UOs are consistent with a bidirectional effect of surprise on dopamine firing, which is an important hallmark of prediction-error signals.[13]

In people with schizophrenia, we did not find any pattern of activity consistent with a prediction-error signal in the whole brain or in the ROI analysis. We did find significant aberrant activity in bilateral vStr, anterior cingulate (BA 32), parietal cortex and prefrontal cortex (BA 10) in people with schizophrenia during expected and unexpected outcomes based on whole-brain analysis (that is, Group × Surprise × Reward). When we ignored the reward omission conditions and tested for activation to URs over ERs, we found group differences in many of the same regions, including parietal cortex, anterior cingulate, and the vStr. In particular, activation after URs in the anterior cingulate was significantly blunted in people

with schizophrenia relative to healthy adults, consistent with other studies implicating this region in error-related feedback and the acquisition of adaptive behavior.[44–46]

Our ROI analysis showed the left ventral striatal response to URs and UR omissions was severely blunted in schizophrenia; whereas, the response to ERs (bilaterally) and EOs (right vStr) was exaggerated. Thus, the pattern of activity in the vStr of people with schizophrenia did not normally differentiate between unexpected and expected events. Previous studies have reported prediction-error signals are attenuated in schizophrenia,[19] exaggerated[20] or even diametrically opposed[18] in the striatum of people with schizophrenia. For instance, Murray et al.[19] reported attenuated prediction-error signals were lateralized in the right vStr, however it is unclear whether the attenuated signal was due to a decrease in responses to URs or an increase in baseline responses to ERs. The interaction of the four conditions across reward and surprise described in the present study demonstrates for the first time that the abnormal prediction-error signal in the vStr is due to aberrant responses to both expected and unexpected events in schizophrenia.

Our novel finding that neural responses to expected events are exaggerated in schizophrenia has implications for theories of psychosis. Gray[4,5] has hypothesized the vStr (nucleus accumbens) acts to receive signals representing the discrepancy or 'mismatch' between actual and predicted events and that hyperdopaminergia in schizophrenia results in repeated 'mismatch' signals when no mismatch has occurred. The exaggerated responses to ERs we clearly found in the right vStr (as well as to a lesser extent in the left vStr) are consistent with the notion of persistent activity to well-predicted stimuli. However, the attenuated neural response to unexpected stimuli in the left vStr of people with schizophrenia indicates a failure to respond when a mismatch has occurred. A failure to respond when a mismatch has occurred is also consistent with an aberrant prediction-error signal, but not in the manner predicted by Gray. Moreover, the persistent BOLD deactivation after EOs we found in the right vStr of people with schizophrenia are not predicted by theories of subcortical hyperdopaminergia. Such theories have either ignored the implications of negative prediction errors[5] or assumed ventral striatal activation is unidirectional.[47] An exaggerated negative BOLD signal implies neural deactivation occurs below baseline in schizophrenia, which is prima facie incompatible with a continuous hyperdopaminergic state.

Despite differences in ventral striatal activity, people with schizophrenia displayed equivalent accuracy to healthy adults in predicting the reward. Extensive animal research demonstrates the vStr is critical for successful reward learning.[48,49] Thus, we might expect performance differences in tasks that require prediction errors to guide behavior. However,

no group differences appeared in our study likely because participants were not required to change or update their expectations for successful performance. In other tasks where participants are required to learn new associations, people with schizophrenia perform worse than healthy adults.[50] Instead, the normal performance in the presence of aberrant neural activity in schizophrenia in the present study indicates prediction-error circuitry was engaged by the task, but not required to successfully complete the task.

While we localized the peak voxels of aberrant activity in the vStr of people with schizophrenia, the cluster extended dorsally to include adjacent regions of the associative striatum. Recent ligand-binding studies have implicated the associative striatum rather than the limbic (ventral) striatum in the pathology of schizophrenia.[51,52] The imprecise nature of fMRI data and the extent of cluster activation does not allow us to exclude the possibility that the pathology may also occur in the associative striatum.

The effect of antipsychotic treatment is an important limitation; however, other studies show aberrant prediction-error-related activity in the striatum in first-episode psychosis,[19] abnormal striatal dopamine function in people who are at risk to develop psychosis[52] and that people with schizophrenia without antipsychotics or in an acute phase of psychosis (and people high in schizotypal personality traits) are more likely to learn from expected outcomes instead of unexpected outcomes,[1] suggesting a primary role of the disease in striatal abnormalities. Here, we found that the exaggerated neural activation to ER in the vStr was positively correlated with symptom severity, further suggesting an influence of disease on brain responses in reward pathways. However, higher doses of antipsychotics were correlated with more vStr deactivation during EOs, suggesting that antipsychotics may be interfering with some aspects of appropriate neural encoding of prediction errors.

The present study found neural responses in the vStr of people with schizophrenia failed to differentiate between expected and unexpected events. The results are consistent with the view of the vStr as a site of prediction-error signaling in the mesolimbic pathway, and clarify the attenuated prediction-error signal in schizophrenia was partially due to persistent responding to expected events as well as blunted responses to unexpected events. The aberrant pattern of neural activation in schizophrenia existed in the presence of antipsychotic treatment, and as such, suggests new therapies are needed to restore normal prediction-error signals in the vStr of people with schizophrenia.

Conflict of interest

The authors declare no conflict of interest.

References

1 Martins Serra A, Jones SH, Toone B, Gray JA. Impaired associative learning in chronic schizophrenics and their first-degree relatives: a study of latent inhibition and the Kamin blocking effect. *Schizophr Res* 2001; **48**: 273–289.

2 Waltz JA, Gold JM. Probabilistic reversal learning impairments in schizophrenia: further evidence of orbitofrontal dysfunction. *Schizophr Res* 2007; **93**: 296–303.

3 Weickert TW, Terrazas A, Bigelow LB, Malley JD, Hyde T, Egan MF et al. Habit and skill learning in schizophrenia: evidence of normal striatal processing with abnormal cortical input. *Learn Mem* 2002; **9**: 430–442.

4 Gray JA. Integrating schizophrenia. *Schizophr Bull* 1998; **24**: 249–266.

5 Gray JA, Feldon J, Rawlins JNP, Smith AD, Hemsley DR. The neuropsychology of schizophrenia. *Behav Brain Sci* 1991; **14**: 1–19.

6 Gold JM, Waltz JA, Prentice KJ, Morris SE, Heerey EA. Reward processing in schizophrenia: a deficit in the representation of value. *Schizophr Bull* 2008; **34**: 835–847.

7 Wise RA. Dopamine and reward: the anhedonia hypothesis 30 years on. *Neurotox Res* 2008; **14**: 169–183.

8 Berns GS, McClure SM, Pagnoni G, Montague PR. Predictability modulates human brain response to reward. *J Neurosci* 2001; **21**: 2793–2798.

9 Rodriguez PF, Aron AR, Poldrack RA. Ventral-striatal/nucleus-accumbens sensitivity to prediction errors during classification learning. *Hum Brain Mapp* 2006; **27**: 306–313.

10 Jensen J, Willeit M, Zipursky RB, Savina I, Smith AJ, Menon M et al. The formation of abnormal associations in schizophrenia: neural and behavioral evidence. *Neuropsychopharmacology* 2008; **33**: 473–479.

11 Kapur S. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am J Psychiatry* 2003; **160**: 13–23.

12 Schultz W, Dickinson A. Neuronal coding of prediction errors. *Annu Rev Neurosci* 2000; **23**: 473–500.

13 Niv Y, Schoenbaum G. Dialogues on prediction errors. *Trends Cogn Sci* 2008; **12**: 265–272.

14 Gray NS, Pickering AD, Hemsley DR, Dawling S, Gray JA. Abolition of latent inhibition by a single 5 mg dose of d-amphetamine in man. *Psychopharmacology (Berl)* 1992; **107**: 425–430.

15 Barch DM, Dowd EC. Goal representations and motivational drive in schizophrenia: the role of prefrontal-striatal interactions. *Schizophr Bull* 2010; **36**: 919–934.

16 Ziauddeen H, Murray GK. The relevance of reward pathways for schizophrenia. *Curr Opin Psychiatry* 2010; **23**: 91–96.

17 Waltz JA, Schweitzer JB, Gold JM, Kurup PK, Ross TJ, Salmeron BJ et al. Patients with schizophrenia have a reduced neural response to both unpredictable and predictable primary reinforcers. *Neuropsychopharmacology* 2009; **34**: 1567–1577.

18 Koch K, Schachtzabel C, Wagner G, Schikora J, Schultz C, Reichenbach JR et al. Altered activation in association with reward-related trial-and-error learning in patients with schizophrenia. *Neuroimage* 2010; **50**: 223–232.

19 Murray GK, Corlett PR, Clark L, Pessiglione M, Blackwell AD, Honey G et al. Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. Mol Psychiatry 2008; 13: 267–276.

20 Walter H, Kammerer H, Frasch K, Spitzer M, Abler B. Altered reward functions in patients on atypical antipsychotic medication in line with the revised dopamine hypothesis of schizophrenia. Psychopharmacology (Berl) 2009; 206: 121–132.

21 First MB, Spitzer RL, Gibbon, Miriam, Williams JBW. Structured Clinical Interview for DSM-IV-TR Axis I Disorders—Patient Edition, 1st revision edn. Biometrics Research Department: New York, NY, 2007.

22 Wechsler D. Wechsler Adult Intelligence Scale, 3rd edn. The Psychological Corporation: San Antonio, TX, 1997.

23 Wechsler D. Wechsler Test of Adult Reading. The Psychological Corporation: San Antonio, TX, 2001.

24 Kay SR, Fiszbein A, Opler LA. The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. Schizophr Bull 1987; 13: 261–276.

25 Nichols T, Hayasaka S. Controlling the familywise error rate in functional neuroimaging: a comparative review. Stat Meth Med Res 2003; 12: 419–446.

26 Genovese CR, Lazar NA, Nichols T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage 2002; 15: 870–878.

27 D'Ardenne K, McClure SM, Nystrom LE, Cohen JD. BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. Science 2008; 319: 1264–1267.

28 Juckel G, Schlagenhauf F, Koslowski M, Filonov D, Wustenberg T, Villringer A et al. Dysfunction of ventral striatal reward prediction in schizophrenic patients treated with typical, not atypical, neuroleptics. Psychopharmacology (Berl) 2006; 187: 222–228.

29 Juckel G, Schlagenhauf F, Koslowski M, Wustenberg T, Villringer A, Knutson B et al. Dysfunction of ventral striatal reward prediction in schizophrenia. Neuroimage 2006; 29: 409–416.

30 Molina V, Hernández JA, Sanz J, Paniagua JC, Hernández AI, Martín C et al. Subcortical and cortical gray matter differences between Kraepelinian and non-Kraepelinian schizophrenia patients identified using voxel-based morphometry. Psychiatry Res 2010; 184: 16–22.

31 Schlagenhauf F, Juckel G, Koslowski M, Kahnt T, Knutson B, Dembler T et al. Reward system activation in schizophrenic patients switched from typical neuroleptics to olanzapine. Psychopharmacology (Berl) 2008; 196: 673–684.

32 Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 2002; 15: 273–289.

33 Mai JK, Paxinos G, Voss T. Atlas of the Human Brain, 3rd edn. Elsevier: Oxford, UK, 2008.

34 Leucht S, Wahlbeck K, Hamann J, Kissling W. New generation antipsychotics versus low-potency conventional antipsychotics: a systematic review and meta-analysis. Lancet 2003; 361: 1581–1589.

35 Woods SW. Chlorpromazine equivalent doses for the newer atypical antipsychotics. J Clin Psychiatry 2003; 64: 663–667.

36 Simon JJ, Biller A, Walther S, Roesch-Ely D, Stippich C, Weisbrod M et al. Neural correlates of reward processing in schizophrenia – relationship to apathy and depression. Schizophrenia Res 2010; 118: 154–161.

37 Rutledge RB, Dean M, Caplin A, Glimcher PW. Testing the reward prediction error hypothesis with an axiomatic model. J Neurosci 2010; 30: 13525–13536.

38 Hare TA, O'Doherty J, Camerer CF, Schultz W, Rangel A. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. J Neurosci 2008; 28: 5623–5630.

39 McClure SM, Berns GS, Montague PR. Temporal prediction errors in a passive learning task activate human striatum. Neuron 2003; 38: 339–346.

40 Spicer J, Galvan A, Hare TA, Voss H, Glover G, Casey B. Sensitivity of the nucleus accumbens to violations in expectation of reward. Neuroimage 2007; 34: 455–461.

41 Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science 1997; 275: 1593–1599.

42 O'Doherty J, Dayan P, Friston K, Critchley H, Dolan RJ. Temporal difference models and reward-related learning in the human brain. Neuron 2003; 38: 329–337.

43 Tobler PN, Dickinson A, Schultz W. Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. J Neurosci 2003; 23: 10402–10410.

44 Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, Cohen JD. Anterior cingulate cortex, error detection, and the online monitoring of performance. Science 1998; 280: 747–749.

45 Holroyd CB, Coles MG. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol Rev 2002; 109: 679–709.

46 Holroyd CB, Nieuwenhuis S, Yeung N, Nystrom L, Mars RB, Coles MG et al. Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. Nat Neurosci 2004; 7: 497–498.

47 Smith A, Li M, Becker S, Kapur S. Dopamine, prediction error and associative learning: a model-based account. Network 2006; 17: 61–84.

48 Iordanova MD. Dopaminergic modulation of appetitive and aversive predictive learning. Rev Neurosci 2009; 20: 383–404.

49 Iordanova MD, McNally GP, Westbrook RF. Opioid receptors in the nucleus accumbens regulate attentional learning in the blocking paradigm. J Neurosci 2006; 26: 4036–4045.

50 Weickert TW, Goldberg TE, Callicott JH, Chen Q, Apud JA, Das S et al. Neural correlates of probabilistic category learning in patients with schizophrenia. J Neurosci 2009; 29: 1244–1254.

51 Kegeles LS, Abi-Dargham A, Frankle WG, Gil R, Cooper TB, Slifstein M et al. Increased synaptic dopamine function in associative regions of the striatum in schizophrenia. Arch Gen Psychiatry 2010; 67: 231–239.

52 Howes OD, Montgomery AJ, Asselin MC, Murray RM, Valli I, Tabraham P et al. Elevated striatal dopamine function linked to prodromal signs of schizophrenia. Arch Gen Psychiatry 2009; 66: 13–20.